

Cursul 1 - INTRODUCERE

- Definiții
 - Colectarea datelor - principii
 - Clasificarea datelor - tipuri de date
 - Prezentarea datelor –
 - Tipuri de tabele
 - Tipuri de grafice
-

Definiție

Statistica este știința

- colectării,
 - clasificării,
 - prezentării,
 - interpretării
datelor și a
folosirii acestora pentru
- ✓ *a formula concluzii și*
 - ✓ *a lua decizii.*
-

Definiții

- **Statistica descriptivă** se ocupă cu colectarea, clasificarea și prezentarea datelor numerice.
 - **Statistica inferențială** (inferential statistics) se ocupă cu interpretarea datelor oferite de statistica descriptivă și cu folosirea acestora pentru a formula concluzii și a lua decizii.
-

Definiții

- **Biostatistica** este știința care se ocupă cu aplicarea metodelor statisticii în științele vieții
 - Vom numi **populație statistică** o mulțime de elemente care au una sau mai multe însușiri comune și care fac obiectul unei cercetări statistice
 - Vom numi **individ statistic** un element al unei populații statistice indiferent de natura acesteia
 - O caracteristică care se schimbă de la individ la individ sau la același individ în timp sau ca răspuns la condițiile de mediu, de boală de medicație, etc, se numește **variabilă**
-

1. Colectarea datelor

PRINCIPIU DE BAZĂ

PRINCIPIUL ATOMICITĂȚII

Informația care se introduce într-o celulă a unui tabel trebuie să nu mai poată fi descompusă în informații mai simple

Ex:

Adresa
Craiova, Brazda lui Novac, Bl A7, Sc II ap 26
Craiova str Lipsyani nr 96a
Craiova, George Enescu, Str Pașcani nr 12

Oras	Cartier	Strada	Nr	Bloc	Scara	Et	Ap
Craiova	Br. Lui Novac			A7	II		26
Craiova		Lipsyani	96a				
Craiova	George Enescu	Pașcani	12				

Greșit!!!

Majoritatea celulelor conțin informație care poate fi separată

CICN	ASC
CICD;AP	Bradycardie sinusala
CICN	Bradycardie sinusala;
CICD;AP	Sechele AVC
CICD;AP	AHC:frate cardiac
CICD;AP	TPSV
CICD	Stenoza carotidiana bilateral
CICD;AP	Bigeminism ventricular
CICD;AP	Bradycardie sinusala
CICN	Hipertiroidie
CICD;AP	Ateromatoza AO; Dislipidemie;Varice m.inf
CICD;AP	BlocAVgr.III; Ex.V ;FO:662
CICD;API	TS; Tiroidectomie pt gusa
CICD;AP	Obezitate
CICN	EPAC; STOP C-R
CICD;AP	BRS major; Bradycardie sinusala
CICD;AP	Hipertrigliceridemie
CICD;AP	Obezitate gr II

Pentru coloana din stânga ar fi corect să creăm câte o coloană separată pentru: CICD, CICN, AP, API, în care să scriem numai **DA** sau **NU**

Introducerea datelor

VSH I	VSH E
60/100	32/54
116/134	
15/35	
28/48	42/64
110/130	36/62
46/75	
5//10	
50/65	35/60
110/135	2//4
10//25	14/30
45/65	35/58
130/140	60/94
40/95	
75/95	
25/40	110/125
30/54	
68/105	
90/100	80/110
120/130	40/60
16/34	
120/140	35/75
90/120	35/65
48/70	
120/135	130/140

Fiecare coloană trebuie înlocuită cu alte două coloane:

VSH I → VSHI1, VSHI2

VSH E → VSHE1, VSHE2

HTA
HTA
HTA IV
HTA
HTA
HTA
HTA II
HTA
HTA II
HTA II
HTA
HTA

Coloana trebuie înlocuită cu două coloane:
HTA, pe care se scrie DA/NU, și
Stadiul, pe care se scrie I / II / III/IV

Designul tabelelor - recomandări

- Să se numereze liniile tabelului, indiferent dacă softul o face automat
 - Să nu se unească mai multe celule pentru a grupa mai multe coloane/linii
 - Să nu se facă câte un tabel separat pentru fiecare categorie mai importantă de pacienți
-

2. Clasificarea datelor

Tipuri de date

- ❑ **Date numerice, întregi sau zecimale** – vârstă, greutate, talie, hemoglobină, transaminaze, calcemie, glicemie, etc
 - ❑ **Date ordinale** – stadiul evolutiv, starea la externare, gradul de rudenie, etc. Codurile folosite au o ordine
 - ❑ **Date nominale** – codul bolii, grupa sanguină, consistența ficatului, culoarea urinei, etc. Codurile folosite **NU** au o ordine
 - ❑ **Date alfanumerice** – nume, prenume, adresă, loc de muncă, descrierea bolii, în general texte sau alte simboluri
-

Tabel de date

NR. CRT.	ANUL	NUMELE	PRENUMELE	SEXUL	MEDIUL	VARSTA	DECADA	PROFESIA	STADIU
1	2008	CALOTA	LUCIA	F	RURAL	62	60-69	FARA OCUPATIE	III
2	2009	CONSTANTIN	MARIN	M	URBAN	55	50-59	FARA OCUPATIE	III
3	2007	FLOREA	ELENA	F	RURAL	83	80-89	PENSIONAR	II
4	2009	HOLT	MARIANA	F	URBAN	65	60-69	PENSIONAR	I
5	2010	IVANESCU	VIRGIL	M	RURAL	64	60-69	PENSIONAR	II
6	2012	LEPADAT	MARIN	M	URBAN	68	60-69	PENSIONAR	III
7	2011	MANOLACHE	EUGENIA	F	RURAL	39	30-39	SALARIAT	IV
8	2010	MARINESCU	DAN	M	RURAL	57	50-59	PENSIONAR	IV
9	2008	STAN	SANDU	M	URBAN	53	50-59	PENSIONAR	V
10	2007	NEAGU	MARIA	M	URBAN	53	50-59	PENSIONAR	III
11	2008	NEDELEA	GHEORGHE	F	RURAL	70	70-79	PENSIONAR	II
12	2009	ORZESCU	ION	M	URBAN	71	70-79	PENSIONAR	V
13	2011	PALIU	MARIN	F	RURAL	76	70-79	PENSIONAR	IV
14	2013	PISICA	MIHAIL	F	RURAL	72	70-79	PENSIONAR	III
15	2010	POPESCU	PETRE	M	URBAN	58	50-59	PENSIONAR	IV
16	2012	PREDA	ION	M	RURAL	45	40-49	SALARIAT	V
17	2009	ALBU	NICOLAE	M	RURAL	45	40-49	SALARIAT	V
18	2008	RADUCAN	ELISABETA	M	URBAN	62	60-69	FARA OCUPATIE	IV
19	2010	RADUCEANU	ION	M	URBAN	39	30-39	FARA OCUPATIE	III
20	2012	IONESCU	MARIA	M	URBAN	39	30-39	FARA OCUPATIE	IV

3. Prezentarea datelor

Tabele de frecvență

Nr.	Clasa	F_i	F_{icc}	F_{icd}	f_i	f_{icc}	f_{icd}
1	25 - 30	5	5	234	2.14%	2.14%	100.00%
2	30 - 35	6	11	229	2.56%	4.70%	97.86%
3	35 - 40	9	20	223	3.85%	8.55%	95.30%
4	40 - 45	26	46	214	11.11%	19.66%	91.45%
5	45 - 50	30	76	188	12.82%	32.48%	80.34%
6	50 - 55	50	126	158	21.37%	53.85%	67.52%
7	55 - 60	53	179	108	22.65%	76.50%	46.15%
8	60 - 65	32	211	55	13.68%	90.17%	23.50%
9	65 - 70	14	225	23	5.98%	96.15%	9.83%
10	70 - 75	5	230	9	2.14%	98.29%	3.85%
11	75 - 80	4	234	4	1.71%	100.00%	1.71%
	Total	234			100%		

Numărul de indivizi =
frecvența absolută

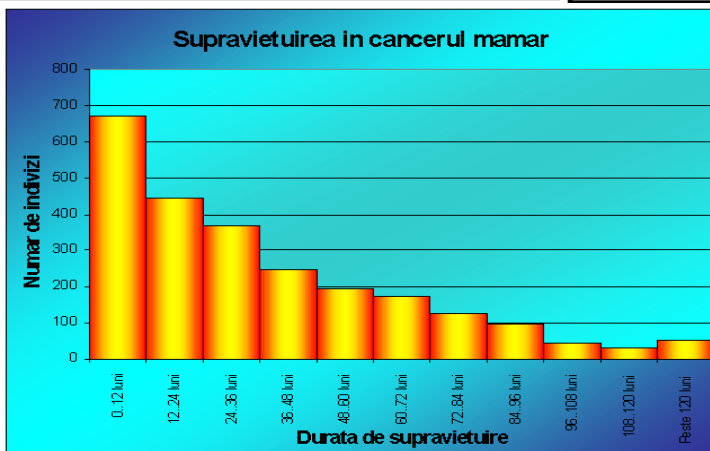
Procentul din total =
frecvența relativă

Tabele de frecvență - exemplu

Supraviețuirea în cancerul de sân

- 2456 pacienți
- Clase de supraviețuire
- Clase din 12 in 12 luni

Nr.Crt	Perioada	Nr.cazuri	Procent %	Procent Cumulat %
1	0..12-luni	672	27.36	27.36
2	12..24-luni	446	18.16	45.52
3	24..36-luni	368	15.00	60.52
4	36..48-luni	249	10.14	70.66
5	48..60-luni	196	8.00	78.66
6	60..72-luni	172	7.00	85.66
7	72..84-luni	126	5.13	90.79
8	84..96-luni	98	4.00	94.79
9	96..108-luni	45	1.83	96.62
10	108..120-luni	31	1.26	97.88
11	Peste-120-luni	52	2.12	100.00



Uneori, datele sunt centralizate ca în tabelul de mai sus.

Reprezentarea datelor sub forma unui grafic (ex: histogramă) le face mult mai ușor de înțeles.

Prezentarea datelor

Tabele de incidență

		Nefropatie		Totali
		"+"	"_"	
Retinopatie	"+"	29 (a)	57 (b)	86
	"_"	2 (c)	172 (d)	174
	Total	31	229	260

Tabelul Retinopatie-Nefropatie (TRN)¶

- ❑ Pacienții cu diabet zaharat pot avea retinopatia și nefropatia ca și complicații majore
- ❑ Spunem că avem o potrivire dacă ambele complicații sunt prezente sau absente
- ❑ În acest tabel sunt 29 potriviri + cu + (celula a) și 172 potriviri – cu – (celula d)

Tabele de incidență

	15 - 19	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	peste 50
sub 15	11	4	3					
15 - 19	194	721	465	60	11			
20 - 24	43	1098	1455	255	78	12		
25 - 29	1	128	1089	457	139	27	5	
30 - 34		3	65	138	115	41	5	
35 - 39	1	1	6	14	65	50	10	2
40 - 44		1		2	7	14	14	7
45 - 49						2	1	1

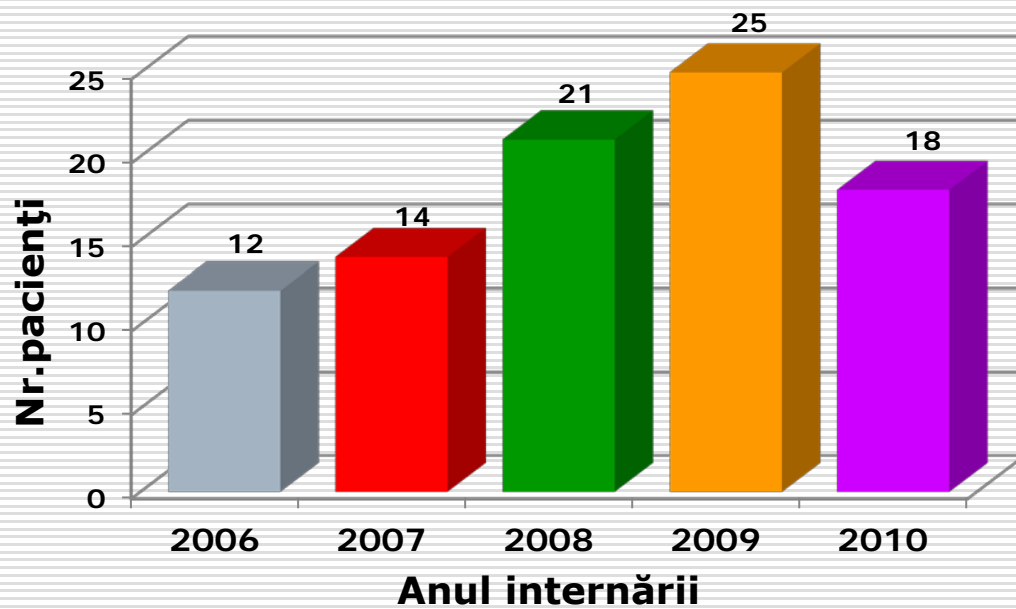
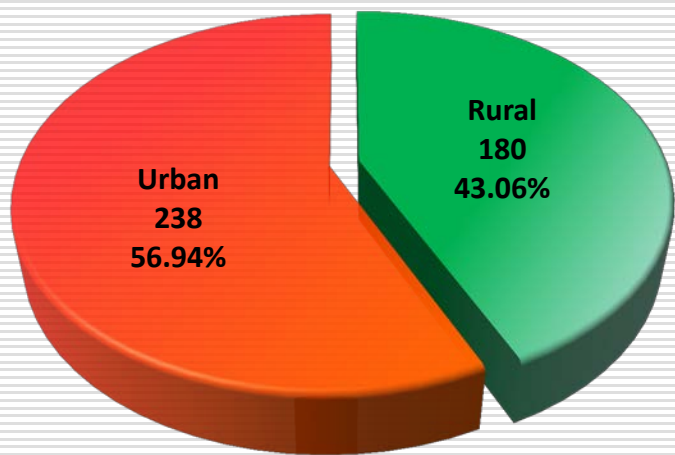
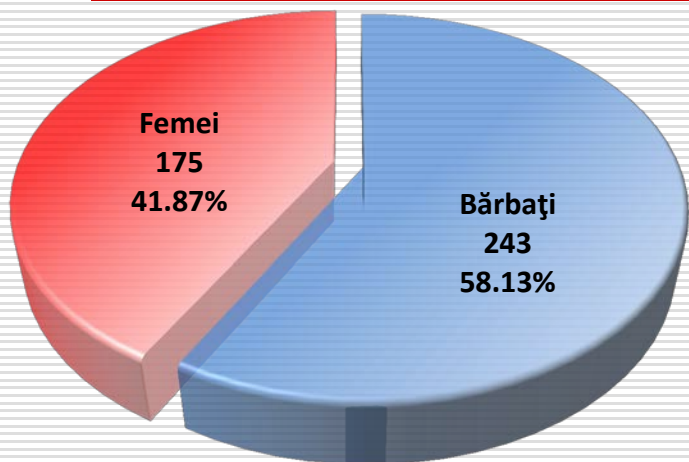
Tabelul de corelație între vârstele mirilor și mireșelor din județul Dolj, din anul 1998.

Pe orizontală, **vârstele mirilor**, iar pe verticală, **vârstele mireșelor**. În fiecare celulă este înscris numărul cuplurilor cu vârstele cuprinse în categoriile respective.

Există o corelație între vârstele mirilor și mireșelor? Evident, DA!

Cele mai multe cazuri se regăsesc pe diagonala principală a tabelului!

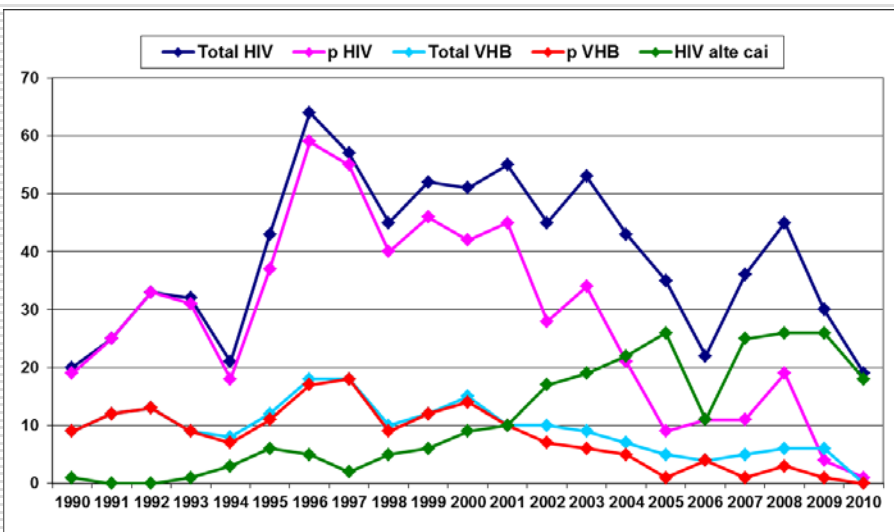
Prezentarea datelor - grafice



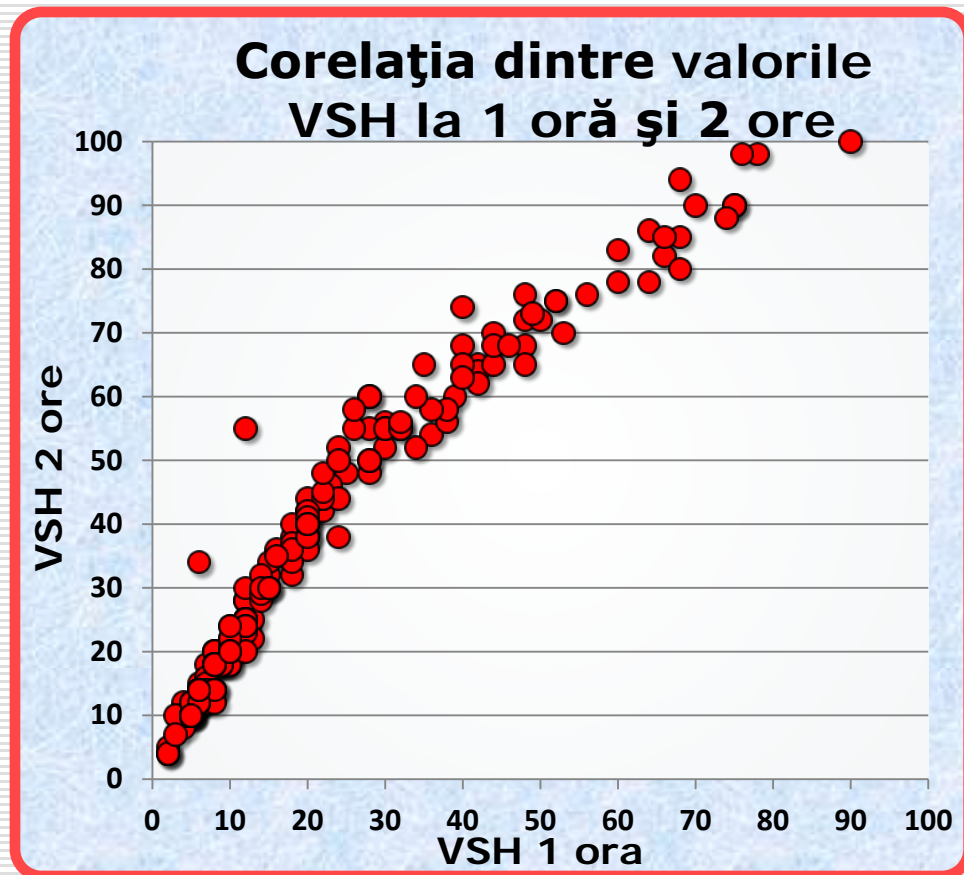
Graficul cu coloane/bare

Graficul circular - Pie

Prezentarea datelor - grafice



**Graficul liniar
(de evoluție)**



Graficul XY (scatter)